OIPE

#2.

RAW SEQUENCE LISTING   DATE: 01/23/2002
PATENT APPLICATION: US/10/040,802   TIME: 19:23:02

Input Set : A:\37036US.txt
Output Set: N:\CRF3\01232002\J040802.raw   **ENTERED**

```
 3 <110> APPLICANT: Saha, K.
 5 <120> TITLE OF INVENTION: Methods and Materials Relating to CD8-Tropic HIV-1
 7 <130> FILE REFERENCE: 28335/37036 US
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/040,802
C--> 9 <141> CURRENT FILING DATE: 2001-12-28
 9 <150> PRIOR APPLICATION NUMBER: US 60/258,472
10 <151> PRIOR FILING DATE: 2000-12-28
12 <160> NUMBER OF SEQ ID NOS: 27
14 <170> SOFTWARE: PatentIn version 3.1
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 2610
18 <212> TYPE: DNA
19 <213> ORGANISM: HIV-1
21 <400> SEQUENCE: 1
22 atgagagtga aggggatcag gaggaattat cagcacttgt ggagatgggg caccatgctc      60
24 cttgggatgt tgatgatctg tagtgctgca gatcaattgt gggtcacagt ctattatggg     120
26 gtgcctgtgt ggaaagaagc aaccaccact ctattttgtg catcagatgc taaagcatat     180
28 agtacagagg tacataatat ttgggccaca catgcctgtg tacccacaga ccccagccca     240
30 caagaaatag taatggaaaa tgtgacagaa gagttcaaca tgtggaaaaa taacatggta     300
32 gaacagatgc atgaggatat aatcagttta tgggatgaaa gcctaaagcc atgtgtaaaa     360
34 ttaactccac tctgtgttac tctaaattgc actactgagt tgaatcttct aaattgcatt     420
36 gataatagta ctaatgataa atgtataccg ccagatcaaa aggagaaat gaaaaactgc      480
38 tctttcaata tcaccgcagg cataagaaat aaggtgcgga agaatatgc acttttttat      540
40 acaagtgatg tagcaccaat agataatgat actatcagtt atagattgat aagttgtaac     600
42 acctcaatca ttacacaggc ctgtccaaag gtatcctttg agccaattcc catacactat     660
44 tgtgccccgg ctggttttgc gattctaaag tgtaaggata ggaatttcaa tggaacagga     720
46 ctatgtaaaa atgtcagcac agtacaatgt acacatggaa ttaggccagt agtatcaact     780
48 caactgctgt aaatggcag tctggcagaa aaagagatag taattagatc tgaaaatttc      840
50 acggacaatg ctaaaaccat aatagtacag ctgcctgaaa tagtacacat taattgtaca     900
52 agacctaaca ataatataag aaaaggtcta cgtataggac cagggagagc atggtggtat     960
54 gcaacaagag gaataatagg aaaaatgaga caaacacatt gcaacattag tagagtaaaa    1020
56 tggaataaca ctttagaaca gatagttaaa aaattaggag acaaatttgg gactaataat    1080
58 aataaaacaa taatatttaa tcaatcctca ggaggggacc cagaaattac aatgcacact    1140
60 tttaattgtg gaggggaatt ttttactgt aatacaacac aactgtttaa tagtacttgg    1200
62 attcggaatg gtactgattg gactcgaaat gatactgaag gatcagacat cactaacgaa    1260
·64 aatatcacgc tcccatgtag aataaaacaa attataaaca tgtggcagaa agtagggaaa    1320
66 gcaatgtatg cccctcccat cagtggacaa attagctgtt cctcaaatat tacagggctg    1380
68 ctattaacac acgatggtgt tgttggtctg tacacggacg cgaacaacgt gaccttcaga    1440
70 ccgggaggag gaaatatgag ggacaattgg agaagtgaat tatataaata taaagtaata    1500
72 aaagttgaac caataggaat agcacccacc aaggcaaaga gaagagtggt gcagagagaa    1560
74 aaaagagcag tgggaatagg agctatgttc cttgggttct tgggaacagc aggaagcgct    1620
76 atgggcgcag cgtcagtgac gctgacggta caagccagac aattattgtc tggtatagtg    1680
78 caacagcaga acaatctgct gagggctatt gaggcgcaac agcatatgtt gcaactcaca    1740
```

RAW SEQUENCE LISTING   DATE: 01/23/2002
PATENT APPLICATION: US/10/040,802   TIME: 19:23:02

        Input Set : A:\37036US.txt
        Output Set: N:\CRF3\01232002\J040802.raw

```
 80 gtctggggca tcaagcagct ccaggcaaga gtcctggctg tggaaagata cctaagggat  1800
 82 caacagctcc tgggaatttg gggttgctct ggaaaactca tttgcaccac tgctgtgcct  1860
 84 tggaatagta gttggagtaa tagatctctg ggagacattt gggaaaatga taacatgacc  1920
 86 tggatgaagt gggaaagaga aattgataat tacacaagct atatatacac cttaattgaa  1980
 88 gaatcgcaga accagcaaga aaagaatgaa ctagaattat tggaattaga caaatgggac  2040
 90 agtttgtgga gttggtttag cataacaaac tggctgtggt atataaaaat attcataatg  2100
 92 atagtaggag gcttgatagg tttaagaata gttttttagtg tgctttctat agtgaataga  2160
 94 gttaggcagg gatactcacc attgtcgttt cagacccggc ccccagcccc gaggggaccc  2220
 96 gacaggcccg aaggaatcga agaagaaggt ggagagagag acagagacag atccggcaga  2280
 98 ttagtggatg gattcttagc acttatctgg gtcgacctgc ggagcctgtg cctcttcagc  2340
100 taccaccgct tgagagactt actcttgatt gcagcgagga ttgtggaact cctgggacgc  2400
102 agggggtggg aagtcctcaa gtattgttgg aatctcctac agtactggag tcaggaacta  2460
104 aagaatagtg ctgttagctt gcttaatacc atagcaatag cagtagctga ggggacagat  2520
106 agggttatag aaataataca aagagcttgt agagctattc tccacatacc tagaagaata  2580
108 agacagggct ttgaaagggc tttgctataa                                     2610
```

111 <210> SEQ ID NO: 2
112 <211> LENGTH: 477
113 <212> TYPE: PRT
114 <213> ORGANISM: HIV-1
116 <400> SEQUENCE: 2

```
117 Met Arg Val Lys Gly Ile Arg Arg Asn Tyr Gln His Leu Trp Arg Trp
118 1               5                   10                  15
120 Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Asp Gln
121             20                  25                  30
124 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
125             35                  40                  45
128 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Ser Thr Glu Val
129     50                  55                  60
132 His Asn Ile Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Ser Pro
133 65                  70                  75                  80
136 Gln Glu Ile Val Met Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys
137             85                  90                  95
140 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
141             100                 105                 110
144 Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
145             115                 120                 125
148 Asn Cys Thr Thr Glu Leu Asn Leu Leu Asn Cys Ile Asp Asn Ser Thr
149     130                 135                 140
152 Asn Asp Lys Cys Ile Pro Pro Asp Gln Lys Gly Glu Met Lys Asn Cys
153 145                 150                 155                 160
156 Ser Phe Asn Ile Thr Ala Gly Ile Arg Asn Lys Val Arg Lys Glu Tyr
157                 165                 170                 175
160 Ala Leu Phe Tyr Thr Ser Asp Val Ala Pro Ile Asp Asn Asp Thr Ile
161             180                 185                 190
164 Ser Tyr Arg Leu Ile Ser Cys Asn Thr Ser Ile Ile Thr Gln Ala Cys
165             195                 200                 205
168 Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala
169     210                 215                 220
172 Gly Phe Ala Ile Leu Lys Cys Lys Asp Arg Asn Phe Asn Gly Thr Gly
```

RAW SEQUENCE LISTING DATE: 01/23/2002
PATENT APPLICATION: US/10/040,802 TIME: 19:23:02

Input Set : A:\37036US.txt
Output Set: N:\CRF3\01232002\J040802.raw

```
173 225                 230                 235                     240
176 Leu Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro
177             245                 250                 255
180 Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu
181             260                 265                 270
184 Ile Val Ile Arg Ser Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile
185         275                 280                 285
188 Val Gln Leu Pro Glu Ile Val His Ile Asn Cys Thr Arg Pro Asn Asn
189     290                 295                 300
192 Asn Ile Arg Lys Gly Leu Arg Ile Gly Pro Gly Arg Ala Trp Trp Tyr
193 305                 310                 315                 320
196 Ala Thr Arg Gly Ile Ile Gly Lys Met Arg Gln Thr His Cys Asn Ile
197             325                 330                 335
200 Ser Arg Val Lys Trp Asn Asn Thr Leu Glu Gln Ile Val Lys Lys Leu
201         340                 345                 350
204 Gly Asp Lys Phe Gly Thr Asn Asn Asn Lys Thr Ile Ile Phe Asn Gln
205     355                 360                 365
208 Ser Ser Gly Gly Asp Pro Glu Ile Thr Met His Thr Phe Asn Cys Gly
209 370                 375                 380
212 Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln Leu Phe Asn Ser Thr Trp
213 385                 390                 395                 400
216 Ile Arg Asn Gly Thr Asp Trp Thr Arg Asn Asp Thr Glu Gly Ser Asp
217             405                 410                 415
220 Ile Thr Asn Glu Asn Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile
221         420                 425                 430
224 Asn Met Trp Gln Lys Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser
225         435                 440                 445
228 Gly Gln Ile Ser Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr His
229     450                 455                 460
232 Asp Gly Val Val Gly Leu Tyr Thr Asp Ala Asn Asn Val
233 465                 470                     475
236 <210> SEQ ID NO: 3
237 <211> LENGTH: 2652
238 <212> TYPE: DNA
239 <213> ORGANISM: HIV-1
241 <400> SEQUENCE: 3
242 atgagagtga agggggatcag gaagaattat cagcacttgt ggagatgggg caccatactc      60
244 cttgggatgt tgatgatctg tagtgctgca gatcaattgt gggtcacagt ctattatggg     120
246 gtgcctgtgt ggaaagaagc aaccaccact ctattttgtg catcagatgc taaagcatat     180
248 agtacagagg tacataatat ttgggccaca catgcctgtg tacccacaga ccccagccca     240
250 caagaaatag taatggaaaa tgtgacagaa gagttcaaca tgtggaaaaa taacatggta     300
252 gaacagatgc atgaggatat aatcagttta tgggatgaaa gcctaaagcc atgtgtaaaa     360
254 ttaactccac tctgtgttac tctaaattgc actactgagt tgaatcttct aaattgcatt     420
256 gataatagta ctagtggtaa taacactgat aatagcacta gtagtaatag tactgatgat     480
258 aaatgtatac cgccagatca aaaaggaaaa atgaaaaact gctctttcaa tatcaccgca     540
260 ggcataagag ataaggtgcg gaaagaatat gcactttttt atacaagtga tgtagcacca     600
262 atagataatg atgctatcag ttatagattg ataagttgta cacctcaat cattacacag      660
264 gcctgtccaa aggtatcctt tgagccaatt cccatacact attgtgcccc ggctggtttt     720
266 gcgattctaa agtgtaagga taggaatttc aatggaacag gactatgtaa aaatgtcagc     780
```

RAW SEQUENCE LISTING   DATE: 01/23/2002
PATENT APPLICATION: US/10/040,802   TIME: 19:23:02

Input Set : A:\37036US.txt
Output Set: N:\CRF3\01232002\J040802.raw

```
268 acagtacaat gtacacatgg aattaggcca gtagtatcaa ctcaactgct gttaaatggc   840
270 agtctggcaa aaaaagagat agtaattaga tctgaaaatt tcacggacaa tgctaaaacc   900
272 ataatagtac agctgcctga aatagtacac attaattgta caagacctaa caataatata   960
274 agaaaaggtc tacgtatagg accagggaga gcatggtggt atgcaacaag aggaataata  1020
276 ggaaaaatga gacaaacaca ttgcaacatt agtagagaaa aatggaataa cactttagaa  1080
278 cagatagtta aaaaattagg agacaaattt gggactaata ataataaaac aataatattt  1140
280 aatcaatcct caggagggga cccagaaatt acaatgcaca cttttaattg tggaggggaa  1200
282 ttttttttact gtaatacaac acaactgttt aatagtactt ggattcggaa tggtactgat  1260
284 tggactcaaa atgatactga aggatcagac atcactaacg aaaatatcac gctcccatgt  1320
286 agaataaaac aaattataaa catgtggcag gaagtaggga aagcaatgta tgcccctccc  1380
288 atcagtggac aaattagctg ttcatcaaat attacagggc tgctattaac acacgatggt  1440
290 gttgttggtc tgtacacgaa cgcgaacaac gtgaccttca gaccgggagg aggaaatatg  1500
292 agggacaatt ggagaagtga attatataaa tataaagtaa taaaagttga accaatagga  1560
294 atagcaccca ccaaggcaaa gagaagagtg gtgcagagag aaaaaagagc agtgggaata  1620
296 ggagctatgt tccttgggtt cttgggaaca gcaggaagcg ctatgggcgc agcgtcagtg  1680
298 acgctgacgg tacaagccag acaattattg tctggtatag tgcaacagca gaacaatctg  1740
300 ctgagggcta ttgaggcgca acagcatatg ttgcaactca cagtctgggg catcaagcag  1800
302 ctccagcaa gagtcctggc tgtggaaaga tacctaaggg atcaacagct cctgggaatt  1860
304 tggggttgct ctggaaaact catttgcacc actgctgtgc cttggaatag tagttggagt  1920
306 aatagatctc tgggagacat ttgggaaaat gataacatga cctggatgaa gtgggaaaga  1980
308 gaaattgata attacacaag ctatatatac accttaattg aagaatcgca gaaccagcaa  2040
310 gaaaagaatg aactagaatt attggaatta gacaaatggg acagtttgtg gagttggttt  2100
312 agcataacaa actggctgtg gtatataaaa atattcataa tgatagtagg aggcttgata  2160
314 ggtttaaaaa tagttttag tgtgctttct atagtgaata gagttaggca gggatactca  2220
316 ccattgtcgt ttcagacccg gcccccagcc ccgaggggac ccgacaggcc cgaaggaatc  2280
318 gaagaagaag gtggagagag agacagagac agatccggca gattagtgga tggattctta  2340
320 gcacttatct gggtcgacct gcggagcctg tgcctcttca gctaccaccg cttgagagac  2400
322 ttactcttga ttgcagcgag gattgtggaa ctcctgggac gcaggggggtg ggaagtcctc  2460
324 aagtattgtt ggaatctcct acagtactgg agtcaggaac taaagaatag tgctgttagc  2520
326 ttgcttaata ccatagcaat agcagtagct gaggggacag atagggttat agaaataata  2580
328 caaagagctt gtagagctat tctccacata cctagaagaa taagacaggg ctttgaaagg  2640
330 gctttgctat aa                                                      2652
333 <210> SEQ ID NO: 4
334 <211> LENGTH: 491
335 <212> TYPE: PRT
336 <213> ORGANISM: HIV-1
338 <400> SEQUENCE: 4
340 Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
341 1               5                  10                  15
344 Gly Thr Ile Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Asp Gln
345                 20              25                  30
348 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
349        35                  40                  45
352 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Ser Thr Glu Val
353     50                  55                  60
356 His Asn Ile Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Ser Pro
357 65                  70                  75                  80
360 Gln Glu Ile Val Met Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys
361                 85                  90                  95
```

RAW SEQUENCE LISTING   DATE: 01/23/2002
PATENT APPLICATION: US/10/040,802   TIME: 19:23:02

Input Set : A:\37036US.txt
Output Set: N:\CRF3\01232002\J040802.raw

```
364 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
365             100             105             110
367 Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
368         115             120             125
371 Asn Cys Thr Thr Glu Leu Asn Leu Leu Asn Cys Ile Asp Asn Ser Thr
372     130             135             140
375 Ser Gly Asn Asn Thr Asp Asn Ser Thr Ser Ser Asn Ser Thr Asp Asp
376 145             150             155             160
379 Lys Cys Ile Pro Pro Asp Gln Lys Gly Lys Met Lys Asn Cys Ser Phe
380             165             170             175
383 Asn Ile Thr Ala Gly Ile Arg Asp Lys Val Arg Lys Glu Tyr Ala Leu
384         180             185             190
387 Phe Tyr Thr Ser Asp Val Ala Pro Ile Asp Asn Asp Ala Ile Ser Tyr
388         195             200             205
391 Arg Leu Ile Ser Cys Asn Thr Ser Ile Ile Thr Gln Ala Cys Pro Lys
392     210             215             220
395 Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
396 225             230             235             240
399 Ala Ile Leu Lys Cys Lys Asp Arg Asn Phe Asn Gly Thr Gly Leu Cys
400             245             250             255
403 Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
404             260             265             270
407 Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Lys Lys Glu Ile Val
408         275             280             285
411 Ile Arg Ser Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln
412     290             295             300
415 Leu Pro Glu Ile Val His Ile Asn Cys Thr Arg Pro Asn Asn Asn Ile
416 305             310             315             320
419 Arg Lys Gly Leu Arg Ile Gly Pro Gly Arg Ala Trp Trp Tyr Ala Thr
420             325             330             335
423 Arg Gly Ile Ile Gly Lys Met Arg Gln Thr His Cys Asn Ile Ser Arg
424             340             345             350
427 Glu Lys Trp Asn Asn Thr Leu Glu Gln Ile Val Lys Lys Leu Gly Asp
428         355             360             365
431 Lys Phe Gly Thr Asn Asn Asn Lys Thr Ile Ile Phe Asn Gln Ser Ser
432     370             375             380
435 Gly Gly Asp Pro Glu Ile Thr Met His Thr Phe Asn Cys Gly Gly Glu
436 385             390             395             400
439 Phe Phe Tyr Cys Asn Thr Thr Gln Leu Phe Asn Ser Thr Trp Ile Arg
440             405             410             415
443 Asn Gly Thr Asp Trp Thr Gln Asn Asp Thr Glu Gly Ser Asp Ile Thr
444             420             425             430
447 Asn Glu Asn Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
448         435             440             445
451 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln
452     450             455             460
455 Ile Ser Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr His Asp Gly
456 465             470             475             480
459 Val Val Gly Leu Tyr Thr Asn Ala Asn Asn Val
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/040,802   DATE: 01/23/2002
                                            TIME: 19:23:03

Input Set : **A:\37036US.txt**
Output Set: **N:\CRF3\01232002\J040802.raw**

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date